Discrete Choice Modeling
William Greene
Stern School of Business, New York University

Lab Session 1 Assignment
Basic Regression and Binary Choice Modeling

I. Basic Regression

This first assignment will help you get started with some familiar, basic estimation and analysis computations. This assignment is based on the German health care data discussed in class. They are an unbalanced panel of 7,293 households observed in 1 to 7 years from 1984 to 1994 (with a couple gaps).

Import the data
File:Import … Import healthcare.csv

For most of our present purposes, however, we will treat theese data as a cross section of 27,326 observations. Since these are panel data, we define them as a panel now – later it will be convenient to move back and forth between panel and pooled data treatments.

xtset id

1. Descriptive Statistics

First, let's take a look at the data. Use

summarize * /// **The * means all variables in the active data set.**

to get some descriptive statistics for the variables in the data set. The analysis below will focus on the income variable. We are going to use log(income) as the dependent variable in the regressions. The following describes this variable and examines whether it appears to be normally distributed by comparing it to a random sample of draws from the normal distribution with the same mean and standard deviation. The kernel estimator uses only the 1994 data.

gen logincome = log(income)
kdensity logincome if year == 1994, normal

The following produces boxplots for the incomes of the female household heads by year. Depending on how they are drawn, boxplots can be distorted by extreme observations in any data set. In the following, we take a simple strategy, and just restrict attention to a range that includes most of the data. The set of plots reveals both the skewed nature of the distribution and the upward trend.

graph box income if female == 1 & income <= 2, over(year)

I am also interested in the education variable. In the original data, education is coded in part years, so a histogram is not very pretty. I will look at the full years of education by converting **educ** to integers.

hist educ if year == 1994

(Note only one graph window can be open at a time.) The distinctive lump at 18 years in the figure probably shows that the data are censored – it appears that education above 18 years is coded as 18. A histogram for a continuous variable will only look good if the data really are continuous. A kernel estimator is not necessarily better. You can see by using the commands

```
gen yearseduc = int(educ)
histogram yearseduc if year == 1994
kdensity yearseduc if year == 1994
```

2. Linear Regression and Testing Hypotheses

For this exercise, we will pool the data and not explicitly use the panel aspect. For convenience, we define a couple of namelists with

```
global demographic age female married
global years year1984 year1985 year1986 year1987 year1988 year1991
```

(We have omitted year1994, so year1994 is the base year.)

To start, we are going to do some linear regression modeling using the variable *logincome* as the dependent variable. We will fit a simple least squares regression with

```
regress logincome $demographic $years
```

An alternative way to handle a categorical variable is to use the internal procedure.

```
regress logincome $demographic i.year
```

(Note that using the internal form changes the base year from 1994 (the last year) to 1984 (the first year).) Of course this is inconsequential, but it does change the normalization of the dummy variable year coefficients. I will want to use the R^2 from this regression, so I save it for later with

```
scalar r20 = e(r2)
```

Does education help to explain the variation in logincome? Add educ to the regression and test the h*ypothesis that the coefficient on education equals zero* using an F test. (We use the familiar Wald (squared t) test first, then construct the F statistic.)

```
regress logincome $demographic i.year educ
test educ  /// Wald statistic
scalar r21 = e(r2)
scalar df1 = e(df_r)
display ((r21 - r20)/1) / ((1 - r21)/df1)  /// F statistic
```

What did you find? Note, your results contain two statistics for carrying out this test, the F statistic and a t statistic reported with the regression results. What are the results?

Now, test the joint hypothesis that neither gender nor education are significant in the model.

```
test educ  female
```

Test the hypothesis that the three coefficients in **demographic** all equal zero. What do you find?

```
test $demographic
```

We also want to test for the presence of 'time' effects in the regression model. In the regression setup, we used i.year to specify the categorical variable. It is necessary to be a little careful here – Stata does not recognize i.year in the test procedure. We can use, instead

```
regr logincome $demographic $years
test $years
```

(More transparent) We can use matrix algebra. The **REGRESS** command provides the coefficients in e(**b**) and covariance matrix in e(V) for us to use in matrix algebra and other commands. For example, in the regression command, the years variables are the 5^{th} to 10^{th} variables. We can use

```
regr logincome $demographic $years educ
matrix vy = e(V)
matrix vy = vy[4..9,4..9]
matrix by = e(b)
matrix by = by[1..1,4..9]
matrix wald = by*invsym(vy)*by'
matrix list wald
```

3. Partial Effects.

Consider the elaborate nonlinear regression model

$$\text{logincome} = \beta_1 + \beta_2\text{age} + \beta_3\text{educ} + \beta_4\text{female} + \beta_5\text{age*educ} + \beta_6\text{age}^2 + \beta_7\text{age*female} + \beta_8\text{educ*female} + \varepsilon$$

What are the partial effects of Age and Educ on logincome? Differentiating, we get

$$\partial\text{logincome}/\partial\text{age} = \beta_2 + \beta_5\text{educ} + 2\beta_6\text{age} + \beta_7\text{female}$$

$$\partial\text{logincome}/\partial\text{educ} = \beta_3 + \beta_5\text{age} + \beta_8\text{female}.$$

What is the male – female income differential?

$$(\text{logincome}|\text{female}=1) - (\text{loginc}|\text{female}=0) = \beta_4 + \beta_7\text{age} + \beta_8\text{educ}.$$

How can you compute these and obtain standard errors for them? There are built in functions that can be used for this sort of computation. First fit the regression with the interaction terms made explicit for the post estimation program.

```
regress logincome c.age c.educ i.female c.age#c.educ c.age#c.age  ///
                  c.age#i.female c.educ#i.female
```

The basic marginal effects (averaged over the sample) are then obtained with

```
margins,dydx(age)
margins,dydx(educ)
margins,dydx(i.female)
```

(a) A more elaborate calculation is the effect of age computed for education fixed at 12,14,16,18,20, and averaged over sample observations.

```
margins,dydx(age) at(educ=(12(2)20))
```

(b) Effect of education computed for ages of 25, 28, 31, …, 64. Plot of the values with confidence intervals.

```
margins,dydx(educ) at(age=(25(3)64))
marginsplot
```

(c) Effect for female, for three levels of education, age 25 to 64 at each education level. Plots of three sets of values. We compute the partial effects and the predictions of the regression.

```
margins,dydx(i.female) at(age=(25(5)65) educ=(12,16,20))
marginsplot
```

A detour (via NLOGIT).
When the model contains a set of categories, such as levels of education, say coded with 4 dummy variables: LTHS (less than high school), HS (high school), COLL (college) or GRAD (postgraduate), the partial effects for each dummy variable compute the effect relative to the base category. It might be interesting to compute the other partial effects. For example, suppose that LTHS is the base. We might compute the impact on income of achieving college education after finishing high school. This suggests a 'transition matrix' of partial effects. In the regression, the educ variable is replaced by the group of variables, in the primary effect and in the interactions.

```
? Examine threshold effects of education
CREATE          ; LTHS    = YrsEduc < 12
                ; HS      = YrsEduc = 12
                ; COLL    = (yrseduc > 12)*(yrseduc<=16)
                ; GRAD    = yrseduc > 16 $
NAMELIST        ; degree  = LTHS,HS,COLL,GRAD $
REGRESS         ; lhs = income
? Note dot after degree. Drops last category when it is expanded.
                ; rhs = one,age,degree., female, degree.*age,
                        age^2, age*female,   degree.*female $
Partials  ; effects:  degree  ;transition $
```

```
---------------------------------------------------------------------------
Ordinary     least squares regression .....
LHS=INCOME   Mean                 =         .35214
             Standard deviation   =         .17686
----------   No. of observations  =          27326  DegFreedom   Mean square
Regression   Sum of Squares       =        86.6414          13       6.66473
Residual     Sum of Squares       =        768.040       27312        .02812
Total        Sum of Squares       =        854.682       27325        .03128
----------   Standard error of e  =         .16769  Root MSE          .16765
Fit          R-squared            =         .10137  R-bar squared     .10095
Model test   F[ 13, 27312]        =      237.00187  Prob F > F*       .00000
--------+------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
  INCOME|  Coefficient       Error       z    |z|>Z*         Interval
--------+------------------------------------------------------------------
Constant|    -.23222***      .02362    -9.83  .0000     -.27851   -.18593
     AGE|     .02883***      .00086    33.44  .0000      .02714    .03052
    LTHS|     .08552***      .01835     4.66  .0000      .04956    .12148
      HS|     .10676***      .02444     4.37  .0000      .05885    .15466
    COLL|     .02968         .02126     1.40  .1627     -.01198    .07135
  FEMALE|     .01132         .01130     1.00  .3166     -.01083    .03346
LTHS*AGE|    -.00562***      .00042   -13.30  .0000     -.00645   -.00479
```

```
  HS*AGE|    -.00458***      .00056    -8.19  .0000     -.00568   -.00349
COLL*AGE|    -.00233***      .00050    -4.67  .0000     -.00330   -.00135
 AGE^2.0|    -.00026***   .8780D-05   -29.80  .0000     -.00028   -.00024
        |Interaction AGE*FEMALE
Intrct05|    -.00089***      .00018    -4.84  .0000     -.00126   -.00053
        |Interaction LTHS*FEMALE
Intrct06|     .02469***      .00878     2.81  .0049      .00748    .04190
        |Interaction HS*FEMALE
Intrct07|     .02505**       .01177     2.13  .0334      .00197    .04812
        |Interaction COLL*FEMALE
Intrct08|     .00910         .01089      .84  .4032     -.01224    .03044
--------+------------------------------------------------------------------
Note: nnnnn.D-xx or D+xx => multiply by 10 to -xx or +xx.
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-------------------------------------------------------------------------------
---------------------------------------------------------------------
Partial Effects  Analysis for Linear Regression Function
---------------------------------------------------------------------
Effects of switches between categories in DEGREE   (dummy variables)
Results are computed by average over sample observations
LTHS    = .7731   HS      = .0632   COLL    = .0950   GRAD    = .0687
---------------------------------------------------------------------
df/dDEGREE         Partial    Standard
From --> To        Effect      Error     |t|  95% Confidence Interval
---------------------------------------------------------------------
LTHS     LTHS       .00000     .00000     .00      .00000      .00000
LTHS     HS         .06668     .00442   15.10      .05802      .07533
LTHS     COLL       .08023     .00386   20.81      .07268      .08779
LTHS     GRAD       .14743     .00444   33.23      .13873      .15613
LTHS     (Other)    .09679     .00261   37.09      .09167      .10190
HS       LTHS      -.06668     .00442   15.10     -.07533     -.05802
HS       HS         .00000     .00000     .00      .00000      .00000
HS       COLL       .01356     .00563    2.41      .00252      .02459
HS       GRAD       .08075     .00604   13.38      .06892      .09258
HS       (Other)   -.04773     .00439   10.86     -.05635     -.03912
COLL     LTHS      -.08023     .00386   20.81     -.08779     -.07268
COLL     HS        -.01356     .00563    2.41     -.02459     -.00252
COLL     COLL       .00000     .00000     .00      .00000      .00000
COLL     GRAD       .06719     .00564   11.90      .05613      .07826
COLL     (Other)   -.06439     .00383   16.79     -.07191     -.05688
GRAD     LTHS      -.14743     .00444   33.23     -.15613     -.13873
GRAD     HS        -.08075     .00604   13.38     -.09258     -.06892
GRAD     COLL      -.06719     .00564   11.90     -.07826     -.05613
GRAD     GRAD       .00000     .00000     .00      .00000      .00000
GRAD     (Other)   -.13471     .00441   30.52     -.14337     -.12606
---------------------------------------------------------------------
* (Other) = conditional share weighted average of all switch effects
---------------------------------------------------------------------

---------------------------------------------------------------
Partial Effects Transition Matrix for DEGREE
There are  4 categories (sample %)
  01=LTHS     (77.31) 02=HS       ( 6.32) 03=COLL     ( 9.50)
  04=GRAD     ( 6.87)
Entry = effect on outcome of switch from row category to column
Switch to Other is unspecified switch out of row category
--------+------------------------------------------------------
        |     01     02     03     04  Other
--------+------------------------------------------------------
    LTHS|   .000   .067   .080   .147   .097
      HS|  -.067   .000   .014   .081  -.048
    COLL|  -.080  -.014   .000   .067  -.064
    GRAD|  -.147  -.081  -.067   .000  -.135
--------+------------------------------------------------------
```

This technique is not available as a built in procedure in Stata. It would be possible to program the computations in part, for example with

gen lths= educ < 12

```
gen hs  =   educ == 12
gen coll =  educ > 12 & educ <= 16
gen grad= educ > 16
global degree lths hs coll grad
gen edlevel = 0*lths + 1*hs + 2*coll + 3*grad
regr logincome c.age i.lths i.hs i.coll i.female c.age#i.lths c.age#i.hs c.age#i.coll ///
c.age#c.age c.age#i.female i.lths#i.female i.hs#i.female i.coll#i.female
margin,dydx(i.hs)
margin,dydx(i.coll)
* Then, the effect we are looking for would be the difference of these two.  It would be necessary
* to use the estimated covariance matrix to compute the standard error for this difference.
```

4. Panel Data

We will examine panel data later in the course. We'll take a brief look at some of the operations here. The GSOEP data are a panel. There is probably correlation across observations, which may mean that although least squares is consistent, the standard errors need correcting. Do we see 'cluster effects' in the standard errors? We consider two approaches. In the first, we correct the OLS standard errors for the correlation across observations in a group. In the second, we use the fixed and random effects approaches to fit the model.

```
global all $demographic $years
regress logincome $all
matrix mols = e(V)
matrix mols = vecdiag(mols)
matrix mols = mols'
regress logincome $all,cluster(id)
matrix mcluster= e(V)
matrix mcluster=vecdiag(mcluster)
matrix mcluster=mcluster'
matrix list mols
matrix list mcluster

xtreg logincome $all, re
xtreg logincome $all, fe
```

Notice the treatment of a time invariant variable in the fixed effects model. Can you see why a second year dummy variable is also dropped from the regression?

II. Binary Choice with Cross Section Data

This exercise will involve estimating and analyzing binary choice models. We will analyze the panel probit, manufacturing innovation data. The data set is PanelProbit.csv. These data are a panel. The data set appears as follows:

```
+-------------------------------------------------------------------------------+
|*************************************************************                  |
|Panel probit data: Stacked, 6350 observations                                  |
|N = 1270, T = 5                                                                |
|EMPLP  = Employment                                                            |
|IM = Industry employment                                                       |
|IP = dependent variable, innovation, binary                                    |
|IMUM = imports share                                                           |
|FDIUM = FDI share                                                              |
|SP = relative size                                                             |
|PROD = productivity                                                            |
|SALES = sales                                                                  |
|LOGSALES = log sales                                                           |
|RAWMTL, INVGOOD, CONSGOOD, FOOD = sector dummies                               |
|T = period, T1,T2,T3,T4,T5 = period dummy variables                            |
|FIRM = firm ID                                                                 |
|Authors Model = (one,logsales,sp,imum,fdium,prod,rawmtl,invgood)               |
|*************************************************************                  |
|Panel Probit Data -  Wide form, 1270 observations                              |
|For observations with T=1 (ignore the others)                                  |
|IP84...IP88 = 5 years of IP                                                    |
|EMPLP84...EMPLP88                                                              |
|IM84...IM88                                                                    |
|IMUM84...IMUM88                                                                |
|FDIUM84...FDIUM88                                                              |
|PROD84...PROD88                                                                |
|SALES84...SALES88                                                              |
|LSALES84...LSALES88                                                            |
+-------------------------------------------------------------------------------+
```

Import the data then declare it to be a panel with

```
xtset firm
```

Some other setup: declare some convenient lists of names.

```
global sector rawmtl invgood consgood
global x im imum fdium sp prod logsales
global allx $x $sector
```

1. Different Functional forms.

As we saw in class, the different distributions chosen for the binary choice model each imply a scaling of the coefficients. Superficially, it appears that the model results depend heavily on the distribution. But, this is illusory. The differences essentially disappear when we examine the partial effects rather than the raw coefficients. The following will illustrate this effect for two specific functional forms.

```
probit ip $x
margin,dydx($x)
logit ip $x
margin,dydx($x)
```

2. The Linear Probability Model

Some recent applications have used linear regression to fit a 'linear probability' model, rather than employ the usual probit or logit model. What does least squares do in a binary choice setting? As might be expected from the previous exercise, the coefficients one obtains are very different. Are the results? The following compares the results of the linear probability model to those of a logit model, both in terms of the coefficients and the partial effects. The results suggest what is actually happening when one uses a linear probability model. The coefficients are approximating the partial effects (at the means of the data) of the appropriate nonlinear binary choice model.

```
regr ip $x
margin,dydx($x)
matrix bols=e(b)
matrix bols = bols'
matrix list bols
probit ip $x
margin,dydx($x)
```

The success of the linear probability at mimicing the probit model is mixed. Notice the good result for IMUM and FDIUM, but the less favorable results for IM, SP, PROD and LOGSALES.

3. A Robust Covariance Matrix.

It is now common to compute a 'robust' sandwich type of estimator when fitting a binary choice model. As we discussed in class, there is not much in the way of failures of the model assumption to which the MLE could be robust. Nonetheless, it might be of interest how much difference it makes. The robust estimator is $\mathbf{H}^{-1}(\mathbf{G}'\mathbf{G})\mathbf{H}^{-1}$, where $\mathbf{H}$ is the negative of the Hessian of the log likelihood and G is the n×K matrix of first derivatives, by observation, of the log densities. The following computes the conventional estimator, $\mathbf{H}^{-1}$ and the robust estimator. We then report the two sets of results.

```
* Conventional estimator
probit ip $x
matrix vmle=e(V)
matrix vmle=vecdiag(vmle)
matrix vmle=vmle'
* 'Robust' estimator
probit ip $x,robust
matrix vrobust=e(V)
matrix vrobust=vecdiag(vrobust)
matrix vrobust=vrobust'
* 'Cluster corrected' estimator
probit ip $x,cluster(firm)
matrix vcluster=e(V)
matrix vcluster=vecdiag(vcluster)
matrix vcluster=vcluster'
matrix list vmle
matrix list vrobust
matrix list vcluster
```

With one notable exception (prod), the so-called robust estimator doesn't matter much. But, the clustering seems to make a large difference. Again, this is to be expected.

4. Creating a Plot of Probabilities.

Once estimation is completed, there are a variety of useful post estimation computations that can be carried out with the estimated model. To begin, it is useful to display the predicted probabilities produced by the model. The following estimates a probit model for innovation, then simulates the probabilities over the range of logSales. The plot is generated by dividing the range into 20 parts from the sample minimum of logSales to the maximum. A listing of the probabilities averaged over the sample with all other variables taking their observed values is shown, followed by a plot with a confidence interval around the prediction.

```
quietly probit ip im imum fdium sp prod logsales if t==1
quietly margins, at(logsales=(3.5(.7)18))
marginsplot
```

5. Fit Measures

The binary choice models are not fit by least squares, and there is no R squared-like statistic to measure the correlation between the predictions of the model and the observed data. Many ad hoc measures have been proposed. The most widely known is McFadden's pseudo R squared, which as discussed in class, does not actually measure anything like the fit of the model to the data. We examined a number of others in class. The following fits a probit model and stores the predicted probabilities. It then computes predictions by the rule 'Predict y = 1 if fitted probability is greater than T*.' The usual choice is T* = .5. You can change the .5 to some other value then see if the value gives a better fit. Some authors label this statistic the 'count R squared,' though that name seems a bit misleading.

```
probit ip $x
predict p
gen iphat = p>.5
table ip iphat
probit
```

6. Partial Effects for a Quadratic and for Interaction Terms

Marginal effects in the binary choice models are complicated functions of the parameters and the data. They are more so when the index function contains complex functions of the data. Suppose, for example,

$$P = \Phi(\boldsymbol{\beta}'\mathbf{x} + \alpha_0 \text{logSales} + \alpha_1 \text{logSales}^2).$$

The marginal effect of logSales, which is the effect on the probability of a one percent change in sales is

$$\partial P/\partial \text{logSales} = \phi(\boldsymbol{\beta}'\mathbf{x} + \alpha_0 \text{logSales} + \alpha_1 \text{logSales}^2) \times (\alpha_0 + 2\alpha_1 \text{logSales})$$

Computing these properly is a longstanding, widely discussed issue in modern software. The problem, in general, is in obtaining the right single effect for logSales rather than separate effects for the two parts, neither of which give the right answer. Recent versions of Stata (with 'Margins') and NLOGIT (with PARTIALS and SIMULATE) have automated the computation of these types of effects. The following does several computations around this formulation. The probit model contains the indicated quadratic term in logSales. The first command computes the average partial effects for logSales and fdium. The second computes the average partial effect for

logSales while varying fdium from .05 to 1.0 in steps of .05, and plots the results. This calculation is done using the delta method.

```
* Partial effects for a nonlinear model
* Partial effects for categories - sectors
global sector $sector food
probit ip $x $sector
margins,dydx($sector)

probit ip im imum fdium sp prod c.logsales c.logsales#c.logsales
margins,dydx(c.logsales)
margins,dyex(fdium)
margins,dydx(c.logsales) at(fdium=(.05(.05)1))
marginsplot
probit ip im c.imum c.fdium c.sp prod c.logsales c.logsales#c.logsales ///
        c.imum#c.sp c.fdium#c.sp
quietly margins,dydx(sp) at (sp=(.05(.05)1.0))
marginsplot
```

7. A Group of Dummy Variables for a Set of Categories

The data set also includes a set of sector dummy variables for four sectors. It might be interesting to examine the different results for the four sectors. The Namelist instruction defines the data matrix Sector which contains all four dummy variables. One of them must be dropped in estimation.

```
* Partial effects for categories - sectors
global sector $sector food
probit ip $x $sector
margins,dydx($sector)
```

8. Testing for Structural Change.

A common test is for homogeneity of the parameter vector across different groups. For example, in our application here, it might be interesting to test whether underlying structural of the model has changed over the five year period of the data. Consider the structure

$$P_{it} = F(\boldsymbol{\beta}_t'\mathbf{x}_{it}),\ i = 1,\ldots,1270,\ t = 1,\ldots,5\ (1993\ \text{to}\ 1997)$$

which allows for different coefficient vectors in each year. We are interested in testing the hypothesis

$$H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta}_3 = \boldsymbol{\beta}_4 = \boldsymbol{\beta}_5$$
$$H_1: \text{ not } H_0.$$

In a linear regression context, this would be a 'Chow' test and would be tested with an F test. Since this is not a linear regression model, we can't use the F test here. The easiest way to do this test is with a likelihood ratio test. The strategy is to fit the restricted model (pool the 5 years of data) and the unrestricted model (estimate the model separately for each year), and compare the log likelihoods. The log likelihood for the unrestricted model is the sum of the five years. Here is how you can automate this computation. Carry out the test. What do you conclude? Should the null hypothesis be rejected? Repeat the test using a logit model instead of a probit model. Does the conclusion change? Try the exercise again while adding the sector dummy variables to the model. To do these, it is only necessary to change the model name from probit to logit, or the global command, **global x im imum fdium sp prod logsales,** by adding variables to it.

```
* Chow style test for structural change
probit ip $x
scalar ll = e(ll)
scalar loglsum = -ll
foreach time in 1 2 3 4 5 {
quietly probit ip $x if t==`time'
scalar loglsum = loglsum + e(ll)
}
display 2*loglsum
```

9. Hypothesis Tests:

This exercise will illustrate two methods of carrying out hypothesis tests. Two tests are carried out. All of the procedures save for the last carry out the test of whether the sector dummy variables should be included in the index function in the probit model. In the last test, The model is

$$y_i^* = \boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i$$
$$\varepsilon \sim N[0,\sigma_i^2],\ \sigma_i = \exp(\boldsymbol{\gamma}'\mathbf{z}_i).$$
$$y_i = 1(y_i^* > 0]$$

and the test of whether $\gamma = 0$ is carried out using an LM test. The (small) advantage of the LM test is that it is not actually necessary to estimate the model to carry out the test as the statistic is based on the restricted, homoscedastic model.

```
* Testing for and estimating a heteroscedastic probit model
probit ip $x if t == 5
scalar logl0 = e(ll)
probit ip $x $sector if t == 5
test $sector
scalar logl1 = e(ll)
display 2*(logl1 - logl0)
hetprob ip $x if t==5,het($sector)
scalar loglh = e(ll)
display 2*(loglh - logl0)
```

10. Simulation:

Using the binary choice model simulator, examine how a 1.1 fold increase in LOGSALES which corresponds to a roughly 10% increase in sales would affect the probability of innovation. The BinaryChoice command carries out a simulated change in every observation, and shows what would happen to the predicted sample responses. The simulation displays the average predicted probabilities over a range of values of logSales.

```
probit ip $x
margins, at(logsales=(5(1)15))
marginsplot
```